云 南 植 物 研 究　2004，**26**（1）：89～95
Acta Botanica Yunnanica

天麻蛋白质的双向电泳和肽质量指纹谱分析与鉴定[*]

张晓勤，胡金勇，曾　英[**]，刘小烛
（中国科学院昆明植物研究所，云南 昆明　650204）

摘要：采用双向聚丙烯酰胺凝胶电泳和质谱技术对天麻染菌球茎皮层和不染菌的新生球茎皮层进行了比较蛋白质组分析与鉴定。双向电泳后在分子量 12～97 kD、等电点 3～10 范围内，每块胶分离到约 900 个蛋白质点。对新生球茎中表达量明显增加的 5 个蛋白质点用基质辅助激光解吸/电离飞行时间质谱（MALDI-TOF MS）进行肽质量指纹谱的分析，并通过检索不同的数据库进行蛋白质鉴定与功能预测，初步认为第 4 号蛋白点是一个与转录有关的 RNA 结合蛋白。同时本文在天麻蛋白质组样品制备、数据库检索策略以及蛋白质鉴定成功率等方面进行了探讨。

关键词：天麻；蛋白质组；双向电泳；肽质量指纹谱
中图分类号：Q 946　　**文献标识码：**A　　**文章编号：**0253－2700(2004)01－0089－07

Identification of Proteins from *Gastrodia elata* by Two-Dimensional Electrophoresis and Peptide Mass Fingerprinting

ZHANG Xiao-Qin，HU Jin-Yong，ZENG Ying[**]，LIU Xiao-Zhu
（*Kunming Institute of Botany*，*Chinese Academy of Sciences*，Kunming 650204，China）

Abstract：Proteins extracted from the cortex of fungi-infected nutritive corms and non-infected secondary corms of *Gastrodia elata* were separated by two dimensional polyacrylamide gel electrophoresis（2-DE）with immobilized pH gradients（IPG）. Reference maps were obtained with a good reproducibility. Approximately 900 proteins per 2-DE gel were isolated with molecular weights ranging 12－97 kD and p*I* 3－10. The spots showed a symmetric shape indicating a good resolution in both electrophoresis dimensions. Five proteins of high levels observed in a secondary corm gel were excised and identified using peptide mass fingerprinting and matrix-assisted laser desorption ionization time-of-flight mass spectrometry. A summary of the identified proteins and their putative functions are presented. One of them，spot no. 4 is identified as a RNA-binding protein involved in transcription. Technical details concerning database queries and protein identification success rates in the absence of a sequenced genome are reported and discussed. Parameters for sample preparation are critically investigated.

Key words：*Gastrodia elata*；Proteome；Two-dimensional electrophoresis；Peptide mass fingerprinting

---

* 基金项目：云南省自然科学基金资助课题（2003C0059M）
** 通讯作者　Author for correspondence
收稿日期：2003－07－23，2003－07－31 接受发表
作者简介：张晓勤（1976－）女，湖北人，硕士研究生，主要从事生物化学研究。

双向电泳（2-DE）是分离蛋白质组分的重要技术，其原理是第一向根据蛋白质的等电点（p*I*）分离蛋白质，第二向根据蛋白质分子量大小达到分离目的。近年来，固相 pH 梯度（IPG）干胶条的应用，克服了两性电解质引起的 pH 不稳定、重复性差及碱性蛋白丢失等问题，因此基于 IPG 技术的双向电泳已成为当前蛋白质组学研究的重要手段之一。

模式植物拟南芥的基因组测序以及重要经济植物水稻基因组框架图的完成，大大推动了植物蛋白质组学的研究，已发表水稻、苜蓿等植物的不同组织器官的蛋白质组（郭奕明等，2002；Watson 等，2003）和植物细胞器如叶绿体、线粒体、质膜等（Prime 等，2000；Bardel 等，2002）的蛋白质组研究。在植物和微生物的相互作用方面，对豆科植物根瘤的蛋白质组研究较为深入（Natera 等，2000），发现一些差异显示的蛋白与固氮作用相关。

本研究的实验材料天麻（*Gastrodia elata*）为真菌寄生型兰科植物，不含叶绿素，其生长发育与蜜环菌（*Arimillaria mellea*）等真菌的关系密切。只有当天麻球茎与蜜环菌建立起营养关系，才能长出粗壮的新生球茎而进一步发育完成其生活史，这些新生球茎不被蜜环菌侵染。研究发现天麻皮层的抗真菌蛋白（GAFP）是天麻新生球茎处于拒菌状态的主要物质，GAFP 在新生球茎的含量是相应染菌球茎的 6 倍（胡忠等，1994）。在天麻的生长发育过程中，除 GAFP 外还有哪些蛋白质的表达量发生变化？本文通过比较天麻染菌球茎（也称为营养茎）与从染菌球茎上长出的不染菌新生球茎（也称为次生球茎）的蛋白质组，分析表达量存在明显差异的蛋白质，并采用肽质量指纹谱（PMF）鉴定差异表达的蛋白质点；同时本文在天麻样品制备、电泳参数的选择、数据库检索策略以及蛋白质鉴定成功率等方面进行了探讨。

1 材料和方法

1.1 实验材料

乌天麻（*Gastrodia elata* Bl. f. *glauca* S. Chow）采自云南昭通，洗净泥沙，给每份天麻个体编号，分别削取每份个体的染菌球茎及从染菌球茎上长出的不染菌新生球茎的表皮及皮层组织，用液氮研磨成粉末，－80℃保存备用。

甲叉双丙烯酰胺、碘乙酰胺为 Fluka 公司产品，丙烯酰胺、甘氨酸、三羟甲基氨基甲烷、四甲基乙二胺、CHAPS、超纯尿素、IPG 缓冲液、IPG 覆盖液、IPG 干胶条 pH3－10（非线性，13 cm 长）购自 Amersham Biosciences 公司。双向电泳标准蛋白质购自 Bio-Rad 公司。

1.2 样品制备

蛋白质成分的提取参照 Tsugita 等（1999）并做如下修改：6 g 材料加入 10 倍体积－20℃ 预冷的含 100 g/L 三氯乙酸及 0.07% 巯基乙醇的丙酮溶液，并于－20℃静置过夜；然后 40 000 × *g*，4 ℃离心 30 min，弃上清；再加入－20℃预冷的含 0.07%巯基乙醇的丙酮溶液，－20℃静置 2 h，离心条件同上；重复上一步，立即离心，弃上清；将沉淀冷冻干燥，干粉于－20 ℃放置备用。称取干粉 0.15 g，加入 2 ml 裂解液（9.5 mol/L 尿素，2% Ampholines，20 g/L CHAPS，10 g/L DTT），30℃水浴保温 30 min，然后 70 000 × *g*，15℃离心 1 h，将上清转移至另一只离心管继续在 70 000 × *g*，15℃离心 30 min，上清液即可用于双向电泳，也可以分装上清液于－80℃放置备用。样品的蛋白质浓度用 Bradford 方法测定（李建武等，1994）。

1.3 双向电泳

第一向固相 pH 梯度等电聚焦，使用 IPG 干胶条（pH 3－10，13 cm，非线性）及 MultiphorⅡ电泳仪。按照 Amersham 公司操作手册（Berkelman 等，1998），稍修改。泡涨胶条过夜，用加样杯上样，样品含 100 μg 蛋白质。等电聚焦程序：300 V，2 h；500 V，2 h；1 000 V，1 h；1 000～3 500 V，3 h；3 500 V，3 h。聚焦完备的 IPG 胶条，先在平衡液Ⅰ（50 mmol/L Tris-HCl，6 mol/L 尿素，30%甘油，40 g/L SDS，痕量溴酚

蓝，使用前加入 10 g/L DTT）中平衡 15 min，再在平衡液 II（DTT 换成 25 g/L 碘乙酰胺，其余组成同 I）中平衡 15 min。转移到 SDS-PAGE 均匀胶上（12.5% T，2.6% C，14 cm×14 cm×0.15 cm），待封胶液凝固后用 Hoefer SE600 Ruby 进行第二向电泳，起初每块胶 10 mA 恒流 15 min，然后每块胶 20 mA 恒流直至溴酚蓝前沿到达玻璃板底部。采用两种方式染色：银染和考马斯亮蓝。银染步骤参考（Yan 等，2000）修改，用 40%乙醇代替甲醇。新生球茎的蛋白质用作质谱分析时，第一向电泳上样 300 μg 蛋白质，电泳胶采用考马斯亮蓝 R-250 染色 1.0 h，脱色过夜。

1.4 图像分析、肽质量指纹谱以及数据库检索

凝胶用 Imagescanner 扫描，按照操作手册并参考相关文献（贾宇峰等，2001）用 Melanie 3.02 软件分析图谱。首先进行点检测，然后选择 5 个分布较好的点标记并配齐（align），接着选择一个参考胶，以重叠方式进行自动匹配。用双向电泳标准蛋白质计算样品胶蛋白质点的分子量和等电点实验值。

选取目的蛋白质点，用解剖刀沿染色边缘切下蛋白点放在少量 Milli-Q 水中，交由上海基康生物技术有限公司用 MALDI-TOF 质谱仪制备肽质量指纹谱。用含 50%乙腈，25 mmol/L 碳酸氢铵溶液将凝胶脱色，胶片真空干燥后用胰蛋白酶水解提取肽段，然后与过量的基质混合，加在样品靶盘上，溶剂挥发后蛋白质与基质形成共结晶，再用脉冲激光照射靶点，导致蛋白质电离，经过质量分析器检测，获得多肽混合物的质量（即肽质量指纹谱），最后与数据库中的数据比较，从而达到鉴定蛋白质的目的，具体方法参考 Courchesne 等（1999）。

将肽质量指纹谱同以下数据库的序列进行比较，在 NCBInr.6.4.2003 和 SwissProt.5.27.2003 中使用程序 MS-Fit（网址为 http：//prospector.ucsf.edu），检索参数为：设定物种拟南芥（*A. thaliana*），匹配的最少肽段数（Minimum Matches）为 4，半胱氨酸未修饰，根据实验值设定较宽的分子量和等电点范围，质量误差（mass tolerance）最低为 100ppm，其余为系统默认值：DNA Frame Translation：6，Considered modifications：|Peptide N-terminal Gln to pyroGlu|Oxidation of M|Protein N-terminus Acetylated|，Max. # Missed Cleavages：1，Peptide N terminus：Hydrogen，Peptide C terminus：Free Acid，Min Parent Ion Matches：1，MOWSE On：1，MOWSE P Factor：0.4，Sort Type：Score Sort。同时还用 PeptIdent（网址为 http：//cn.expasy.org/）检索数据库 Swiss-Prot and TrEMBL，物种设定为植物界（VIRIDIPLANTAE），其余参数同上。蛋白质功能分析在蛋白质功能数据库 Pfam（http：//www.sanger.ac.uk/Software/Pfam/）中进行。

2 结果和讨论

2.1 样品制备与双向电泳

天麻样品含大量多糖，若样品制备时直接用裂解液裂解，离心转速为 15 000×*g*，则电泳胶上的蛋白质点稀少（约 50 个），而且横向拖尾严重；采用三氯乙酸－丙酮沉淀后再裂解，并将制样时的离心转速提高到 70 000×*g*，去除了大部分多糖杂质，分离到的蛋白质组分丰富许多，每个电泳胶可得到约 900 个蛋白质点，这些蛋白质点大多集中于 14～66 kD。此外，制样时的离心温度也很关键，有文献报道离心温度为 4℃，但是在实验中发现 4℃离心后管壁有许多尿素结晶析出，因此本实验将离心温度提高到 15℃，不再出现结晶。

经分析，天麻球茎蛋白质双向电泳图谱的重复性较高，不同个体之间也有较好的重复性。在分子量 12～97 kD、等电点 3～10 范围内，共计分离到约 900 个蛋白质点（图 1）。由于天麻营养茎已被蜜环菌感染，其蛋白质组必然包含蜜环菌的蛋白质成分，而要从天麻皮层将正在被消化分解的菌丝单独分离出来提取蛋白质几乎不太可能。我们已成功制备离体培养菌丝的蛋白质双向电泳图谱（未发表），但考虑到离体培养菌丝的蛋白质组并不能完全反映天麻皮层中正在被消化的蜜环菌蛋白质组，且营养茎中看似增量表达的蛋白点可

能含有蜜环菌蛋白质。因此从新生球茎的双向电泳胶上选取5个表达量明显高于染菌球茎的蛋白质点（图1，箭头标记1~5），这5个蛋白质点在不同的天麻个体表达量相当，都表现为新生球茎中的含量明显高于其相应的染菌球茎。

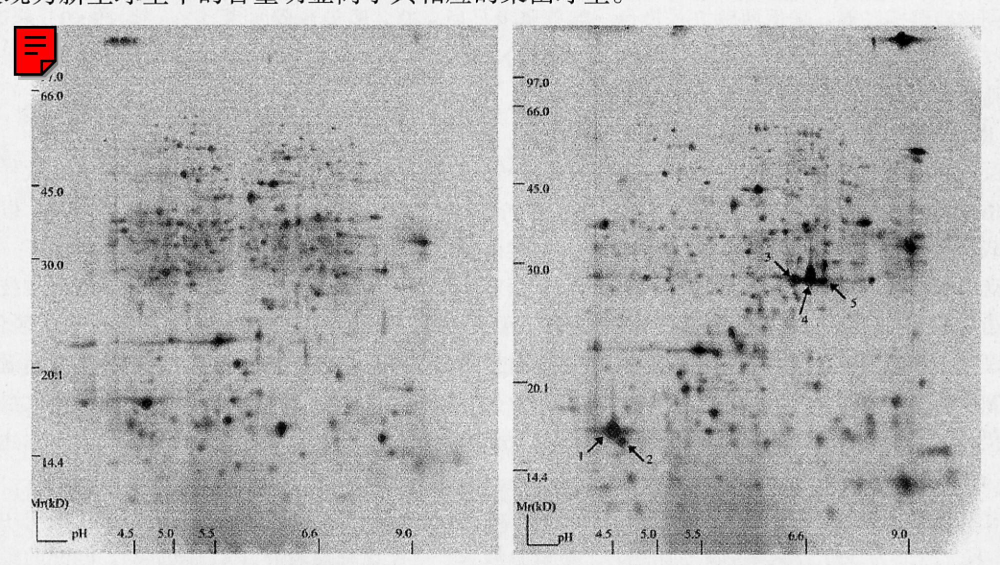


图1 天麻营养茎（左）和次生茎（右）双向电泳银染图谱

箭头和数字标记的蛋白点用于质谱鉴定，图中的蛋白编号与表1中列出的蛋白点编号一致，蛋白质上样量为100 μg，IPG胶条为非线性，13 cm长，pH 3~10，第二向SDS-PAGE凝胶浓度为12.5%（w/v），凝胶银染与脱色参照实验方法。

Fig. 1 2-DE reference maps obtained for nutritive corm (left) and secondary corm (right)

Proteins that identified in this study are marked with arrows and numbers. The numbers correlate with protein identifications listed in Table 1. 2-DE was performed using 100 μg of protein, nonlinear 13 cm IPG strips (pH 3 – 10), and a 12.5% (w/v) total acrylamide SDS second dimension. Gels were silver stained, destained according to the method and images recorded.

2.2 数据库检索与蛋白质鉴定

2.2.1 数据库的检索策略 根据质谱分析获得每个样品的一组同位素峰值，设定物种或物种门类，在较宽泛的分子量和等电点范围内，逐渐升高质量误差，采用MS-Fit和Pept-Ident分别进行检索，结果见表1。表1综合了对天麻新生球茎5个蛋白质点进行分析鉴定的基础数据和初步结果，包括蛋白质丰度、分子量和等电点实验值、检索的数据库以及检索时的质量误差、匹配的肽段数、肽序列对匹配蛋白的序列覆盖率、理论分子量和等电点、匹配蛋白的名称及其GenBank登记号、对应的物种、蛋白质功能。

设定的质量误差较低，检索结果的可信度就较高。但本次检索中如果将质量误差设定为100ppm，几乎都没有匹配的条目；将质量误差设定在200ppm或0.25~1.0Da，检索成功率增加。由于待鉴定的蛋白质来源于兰科植物，而到目前为止还没有公布任何兰科植物的基因组全序列信息，因此将物种范围设定在植物界或拟南芥进行检索，可提高基于序列相似性匹配蛋白质的可信度。

在豆科植物苜蓿各器官的蛋白质组研究中（Watson等，2003），发现仅在蛋白质数据库（SwissProt和NCBInr）进行检索，蛋白质鉴定的成功率只有25%；而在EST库平均成功率为46%。当然，那些不含poly（A^+）尾巴的mRNAs（如线粒体和叶绿体编码的mRNAs）

表1　天麻蛋白质鉴定的相关数据

Table 1　Summary of identified *Gastrodia elata* proteins with supporting data

Spot #	Relative Abund. /%	Exp. MW (kD) /p*I*	Databases/Mass tolerance	# Peptides	Pr Cov /%	Theoret MW (kD) /p*I*	Identification/Comment	Accession Number	Organism Matched	Function
1	3.170	17/4.6	N，1.0Da	5	23.0	28/5.4	MA3domain-containing protein	NP_194797	*A. thaliana*	Disease/defense
				4	19.0	17/5.5	nucleoside-diphosphate kinase	T51612	*A. thaliana*	Metabolism
2	0.884	16/4.7	N，0.5Da	4	19.0	19/5.5	glycosyltransferase family 28	NP_193404	*A. sthaliana*	Metabolism
			S/T，1.0Da	5	29.7	21/5.6	lactoylglutathione lyase	O04885	*B. juncea*	Metabolism
3	1.714	28/6.2	N/S，200ppm	4	12.0	40/5.2	leucoanthocyanidin dioxygenase	Q96323	*A. thaliana*	Secondary metabolism
			S/T，0.5Da	7	19.4	40/5.2	leucoanthocyanidin dioxygenase	Q96323	*A. thaliana*	Secondary metabolism
4	2.643	29/6.4	N，0.6Da	5	31.0	30/6.7	Probable RNA-binding protein	D96792	*A. thaliana*	Transcription
				4	30.0	26/7.0	Glutathione transferase，putative	NP_180505	*A. thaliana*	Metabolism
5	0.836	28/6.5	N，200ppm	6	22.0	32/5.4	expressed protein	NP_195060	*A. thaliana*	Unclear

注：注册登记号来自 GenBank，N = NCBInr.，S = SwissPro，S/T = Swiss-Prot and TrEMB。**检索数据库** N **或** S **时用** MS-Fit **程序；**S/T **用** PeptIdent **程序且限制在植物界。表中包含蛋白点的编号（图1），每个蛋白点的相对丰度，分子量和等电点的实验值，检索的数据库，质谱误差，相匹配的肽段数目，蛋白质序列覆盖率，分子量和等电点的理论值，最佳匹配的数据库登记号，匹配蛋白的生物来源以及蛋白质的功能。**

Note：Accession number is GenBank number.　N = NCBInr. 6.4.2003（1441081 entries），S = SwissPro. 5.27.2003（126844 entries），S/T = Swiss-Prot and TrEMBL（Swiss-Prot Release 41.11 of 06-Jun-2003：127863 entries；TrEMBL Release 23.15 of 06-Jun-2003：857950 entries），MS-Fit was used for searching database N or S，PeptIdent was performed within VIRIDIPLANTAE of Swiss-Prot and TrEMBL. These data include an assigned protein spot number（see Figure 1），percent relative abundance of each protein，experimental M_r and p*I*，database that yield concurrent identifications，mass tolerance，number of peptides matched，percent protein coverage，theoretical M_r and p*I*，database accession number of the best match，the organism to which the matching protein was identified through similarity，and protein function.

很难出现在 EST 库，因此同时检索多种数据库可以提高蛋白质鉴定的成功率。

2.2.2 蛋白质的鉴定 用 PMF 鉴定蛋白质主要基于以下参数：匹配的肽段数、质量精度（*m*/*z* accuracy）、肽序列对匹配蛋白全长的序列覆盖率（最低 10%）、肽谱的质量、匹配峰的强度（最小值 18%～20%）、分子量和等电点实验值与理论值的差异、匹配蛋白所属的物种（Watson 等，2003）。蛋白质的功能分析在蛋白质功能数据库 Pfam（http://www.sanger.ac.uk/Software/Pfam/）中进行；按照 Bevan 等（1998）发表在 Nature 上的植物基因功能分类方案，划分匹配蛋白质的功能类别。

分析比较数据库的检索结果，筛选最适匹配的记录，发现用两种方式都被鉴定为同一种蛋白质的只有第 3 号蛋白（spot no.3）— 一种花青苷合成酶，而第 5 号蛋白（spot no.5）与拟南芥的一种未知功能的表达蛋白最相似。以第 4 号蛋白点（spot no.4）为例，其肽质量指纹谱见图 2，相关的检索记录中以 RNA 结合蛋白（Acc.#：D96792）匹配最好，检索结果见表 2。根据植物基因的功能分类方案，RNA 结合蛋白的功能类别是转录，它在天麻新生球茎大量表达（丰度为 2.6%），可能与旺盛的生长发育密切相关。

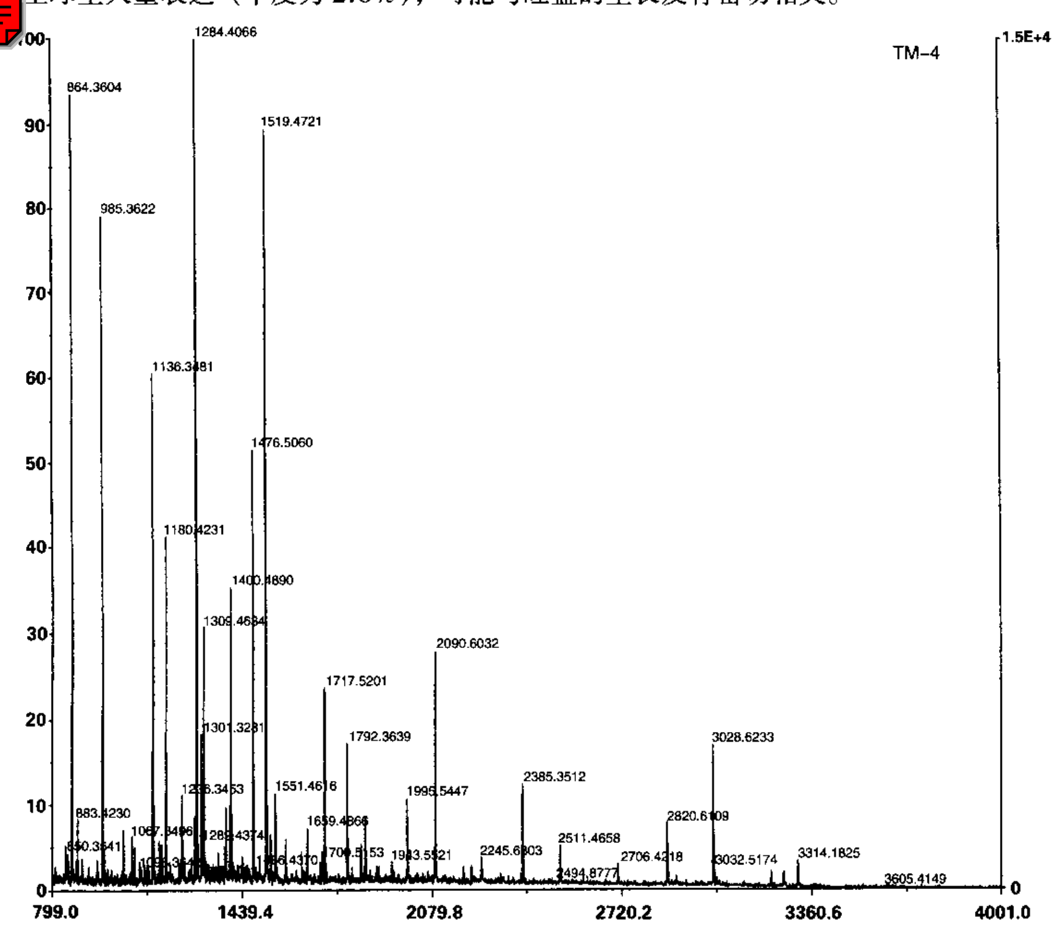


图 2 双向电泳图上第 4 号蛋白点的肽质量指纹谱

Fig. 2 PMF of spot no.4 (TM-4) from 2-DE gel

表 2　第 4 号蛋白点数据库检索结果（登记号 #：D96792）

Table 2　Database querying result of spot no. 4（Acc. #：D96792）

m/z Submitted	MH+ Matched	Delta Da	Missed Cleavages	Peptide
1284.4070	1284.6701	－0.26	0	ISNHFHITTSK
1384.4630	1384.7007	－0.24	1	ACADPTPIIDGRR
1384.4630	1384.7007	－0.24	1	RACADPTPIIDGR
1792.3640	1792.9234	－0.56	0	VFVGGLAWETQSETLR
2089.5970	2089.0606	0.54	0	QHFEQYGEILEAVVIADK
3312.2590	3311.7985	0.46	1	RANCNLASLGRPRPPLPYAVIPNMPDLLPR

总的来看，本文 5 个蛋白点的肽序列覆盖率都没有超过 50%，不是同一物种来源可能是原因之一，如果样品来源于基因组全序列测序完成的物种，相信用 PMF 鉴定蛋白质的成功率和准确度都会增加。对于人类蛋白质组研究，其蛋白质鉴定成功率在 60% 左右（Watson 等，2003）。不同的检索方法得出的检索结果不一致，若要准确鉴定这些蛋白质还应该进一步做串联质谱，PMF 和肽序列联合检索将极大提高蛋白质鉴定的成功率（Watson 等，2003）。

〔参　考　文　献〕

李建武，萧能庚，余瑞元等，1994. 生物化学实验原理和方法［M］. 北京：北京大学出版社，174—176

Bardel J，Louwagie M，Jaquinod M，*et al*，2002. A survey of the plant mitochondrial proteome in relation to development［J］. *Proteomics*，**2**（7）：880—898

Berkelman T，Stenstedt T，1998. 2-D Electrophoresis Using Immobilized pH Gradients Principles and Methods［M］. Sweden：Amersham Pharmacia Biotech，1—47

Bevan M，Bancroft I，Bent E，*et al*，1998. Analysis of 1.9Mb of contiguous sequence from chromosome 4 of *Arabidopsis thaliana*［J］. *Nature*，**391**（6666）：485—488

Courchesne PL，Patterson SD，1999. Idendification of proteins by matrix-assisted laser desorption/ionization mass spectrometry using peptide and fragment ion masses［A］. In：Link AJ ed. 2-D Proteome Analysis Protocols［C］，New Jersey，USA：Humana Press，487—511

Guo YM（郭奕明），Shen SH（沈世华），Jing YX（荆玉祥），*et al*，2002. Plant proteomics in the post-era［J］. *Acta Bot Sin*（植物学报），**44**（6）：631—641

Hu Z（胡忠），Huang QZ（黄清藻），1994. Induction and accumulation of the antifungal protein in *Gastrodia elata*［J］. *Acta Bot Yunnan*（云南植物研究），**16**（2）：169—177

Jia YF（贾宇峰），Lin QX（林秋霞），Guo YJ（郭尧君），*et al*，2001. The image analysis of two-dimensional gel electrophoresis［J］. *Prog Biochem Biophys*（生物化学与生物物理进展），**28**（2）：246—250

Natera SH，Guerreiro N，Djordjevic MA，2000. Proteome analysis of differentially displayed proteins as a tool for the investigation of symbiosis［J］. *Molecular Plant-Microbe Interaction*，**13**（10）：995—1009

Prime TA，Sherrier DJ，Mahon P，*et al*，2000. A proteomics analysis of organelles from *Arabidopsis thaliana*［J］. *Electrophoresis*，**21**（16）：3488—3499

Tsugita A，Kamo M，1999. 2-D electrophoresis of plant proteins［A］. In：Link AJ ed. 2-D Proteome Analysis Protocols［C］，New Jersey，USA：Humana Press，95—97

Watson BS，Asirvatham VS，Wang LJ，*et al*，2003. Mapping the proteome of barrel medic（*Medicago truncatula*）［J］. *Plant Physiol*，**131**（3）：1104—1123

Yan JX，Wait R，Berkelman T，*et al*，2000. A modified silver staining protocol for visualization of proteins compatible with matrix-assisted laser desorption/ionization and electrospray ionization-mass spectrometry［J］. *Electrophoresis*，**21**（17）：3666—3672